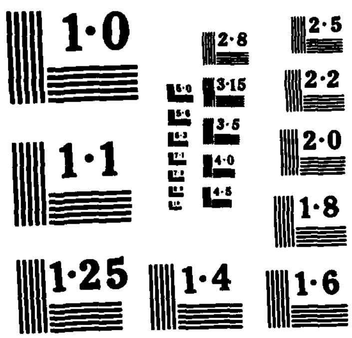
AD-A158 151 B-SPLINES AND LINEAR COMBINATIONS OF UNIFORM ORDER STATISTICS(U) WISCONSIN UNIV-MADISON MATHEMATICS RESEARCH CENTER A G IGNATOV ET AL. MAY 85 MRC-TSR-2817 1/1

UNCLASSIFIED DAAG29-80-C-0041 F/G 12/1 NL

END

FILMED

DTIC




NATIONAL BUREAU OF STANDARDS
MICROCOPY RESOLUTION TEST CHART

AD-A158 151

MRC Technical Summary Report #2817

# B-SPLINES AND LINEAR COMBINATIONS OF UNIFORM ORDER STATISTICS

Z. G. Ignatov and V. K. Kaishev

**Mathematics Research Center**
**University of Wisconsin—Madison**
**610 Walnut Street**
**Madison, Wisconsin 53705**

May 1985

(Received December 14, 1984)

DTIC FILE COPY

**Approved for public release**
**Distribution unlimited**

Sponsored by

U. S. Army Research Office
P. O. Box 12211
Research Triangle Park
North Carolina 27709

UNIVERSITY OF WISCONSIN-MADISON
MATHEMATICS RESEARCH CENTER

# B-SPLINES AND LINEAR COMBINATIONS OF UNIFORM ORDER STATISTICS

Z. G. Ignatov and V. K. Kaishev*

Technical Summary Report #2817

May 1985

## ABSTRACT

Consider independent random variables $X_1, X_2, \ldots, X_n$, uniformly distributed on unequal intervals $(0, a_i)$, $a_i$ real, $i = 1,2,\ldots,n$. Let $S = X_1 + X_2 + \ldots + X_n$. The density of $S$ is shown to be a linear combination of $n!$ B-splines with constant coefficients. A probabilistic interpretation of the B-spline as the density of a linear combination of order statistics from the uniform distribution on $(0,1)$ is given. This interpretation makes it possible to establish recurrence relations for densities and moments of linear combinations of order statistics and to give asymptotic results both for B-splines and linear combinations as well. Examples and applications are also discussed.

AMS 1980 Subject Classifications: Primary 62G30; Secondary 60E05

Key Words and Phrases: B-spline, discrete B-spline, linear combination of order statistics, Polya distribution.

Work Unit Number 4 (Statistics and Probability)

---

*This work was initiated at the Institute of Mathematics, Bulgarian Academy of Sciences and completed while the second author was a participant in the scientific exchange program between the National Academy of Sciences of United States and Bulgarian Academy of Sciences.

Sponsored by the United States Army under Contract No. DAAG29-80-C-0041.

## SIGNIFICANCE AND EXPLANATION

B-splines and linear combinations of order statistics play an important role in the approximation of functions and statistics respectively. It is useful to investigate the relation between the two subjects, both from the theoretical and computational point of view.

B-splines on equidistant knots were viewed nondeterministically by I. J. Schoenberg already in his first work on B-splines. Several studies on distributions of sums of independent uniformly distributed random variables have given formulae which in fact are easily seen to represent splines. However, a precise probabilistic interpretation of the B-spline seems not to be explicitly stated in the literature.

It is shown ~~here~~ in this document that the density of the sum of $n$ independent random variables uniformly distributed on unequal intervals is given by a linear combination of $n!$ B-splines with constant coefficients.

Another useful representation of the same density is given using de Boor's definition of the discrete B-spline.

It is also shown that the B-spline is the density of a linear combination of order statistics from the uniform distribution on $(0,1)$. This interpretation of the B-spline allows one to establish easily validity of recurrence relations for densities and moments of linear combinations of order statistics as well as to state limit theorems both for B-splines and linear combinations.

Examples given below illustrate how asymptotic results for B-splines may be applied to linear combinations of order statistics. Using limit theorems of probability theory two examples of Curry and Schoenberg (1966) for B-splines are derived even in somewhat greater generality (see example 3).

---

The responsibility for the wording and views expressed in this descriptive summary lies with MRC, and not with the author of this report.

# B-SPLINES AND LINEAR COMBINATIONS OF UNIFORM ORDER STATISTICS

Z. G. Ignatov and V. K. Kaishev*

## 1. Introduction and summary.

Let $X_1, X_2, \ldots, X_n$ be independent random variables and $X_i$ be uniformly distributed on the interval $(0, a_i)$, $a_i > 0$, $i = 1, 2, \ldots, n$. Let $S = X_1 + X_2 + \ldots + X_n$. An expression for the density of $S$ was found by Olds (1952), applying the convolution formula for the density of a sum of two random variables and an induction argument. Dempster and Kleyle (1968) have given a geometric derivation of the distribution of $S$ under the condition $\sum_{i=1}^{n} X_i/a_i < 1$ and have pointed out the relation of the latter to the distribution of linear combinations of uniform order statistics on $(0,1)$. The distribution of such linear combinations of order statistics in more generality, covering the case of zero terms in the linear combination has been obtained by Ali (1969) and Weisberg (1971). Cicchitelli (1976) approached the same distributional problem by direct integration. A related distribution was recently considered by Currie (1981).

In the present note the relation of the above mentioned distribution to polynomial splines is emphasized. Theorem 2.1, Section 2 gives an exact representation of the density of $S$ as a linear combination of $n!$ B-splines. The B-spline $M_n(x; x_0, x_1, \ldots, x_n)$ introduced by Curry and Schoenberg (1966), of order $n$, (degree $n - 1$) with knots $x_0 < x_1 < \ldots < x_n$ is defined by the familiar divided difference representation

$$M_n(x; x_0, x_1, \ldots, x_n) = n \sum_{j=0}^{n} \frac{(x_j - x)_+^{n-1}}{\omega'(x_j)} , \tag{1.1}$$

where $\omega(x) = (x - x_0)(x - x_1) \ldots (x - x_n)$, $(x)_+ = \max\{x, 0\}$.

*This work was initiated at the Institute of Mathematics, Bulgarian Academy of Sciences and completed while the second author was a participant in the scientific exchange program between the National Academy of Sciences of United States and Bulgarian Academy of Sciences.

Sponsored by the United States Army under Contract No. DAAG29-80-C-0041.

Let $Z_1, Z_2, \ldots, Z_n$ be independent uniformly distributed random variables on (0,1). Let $Z_{(1)} \geq Z_{(2)} \geq \ldots \geq Z_{(n)}$ be the order statistics of $Z_1, Z_2, \ldots, Z_n$. It is shown in Section 3, Lemma 3.1 that the B-spline coincides with the density of the linear combination $L = a_1 Z_{(1)} + a_2 Z_{(2)} + \ldots + a_n Z_{(n)}$. Recurrence relations for densities of linear combinations of order statistics and their moments are also given. In particular, Lemma 3.2 of Section 3 gives a probabilistic analogue of the well known recurrent relation for the univariate B-splines (c.f. de Boor (1972), Cox (1972)) in terms of densities of linear combinations of order statistics. On the basis of Lemma 3.1 and the limit theorems 6 and 7 Section 5, of Curry and Schoenberg (1966) for the B-spline, two limit theorems for the linear combination of the order statistics $Z_{(1)}, Z_{(2)}, \ldots, Z_{(n)}$ are stated (Section 4, Theorems 4.1, 4.2). Theorem 4.3 of Section 4 gives a necessary and suficient condition for convergence of the B-spline to $1/\sqrt{2\pi}\, \exp(-x^2/2)$ when $n$ tends to infinity. Under this condition, Theorem 4.4 states an expansion for $\int_{-\infty}^{x} M_n(t; x_0, \ldots, x_n)dt$. Both theorems 4.3 and 4.4 are easily obtained by correspondingly applying Lemma 3.1 to the results of Hecker (1976) and van Zwet (1979) for linear combinations of uniform order statistics.

In Section 5, examples and applications of the results of Section 3 and 4 are given. They illustrate both the stochastic approach to studying asymptotic behavior of fundamental spline-functions, and the use of spline theory techniques in probabilistic problems.

## 2. B-splines and the distribution of S.

We first state the following Lemma which interprets the B-spline stochastically.

**LEMMA 2.1.**

The conditional density function $f_S(x/X_1/a_1 > X_2/a_2 > \dots > X_n/a_n)$ coincides $\forall x$ with the B-spline $M_n(x;x_0,x_1,\dots,x_n)$ having knots $x_0 = 0$, $x_1 = a_1$, $x_2 = a_1 + a_2,\dots,$ $x_n = a_1 + a_2 + \dots + a_n$.

**PROOF.**

As assumed above the random vector $(X_1,X_2,\dots,X_n)$ is uniformly distributed in the n-dimensional hypercube $(0,a_1) \times (0,a_2) \times \dots \times (0,a_n) = \Omega$. The random event $\{X_1/a_1 > X_2/a_2 > \dots > X_n/a_n\} = \sigma \subset \Omega$ is an n-simplex denoted by $\sigma$ with vertices $(0,0,\dots,0)$, $(a_1,0,0,\dots,0)$, $(a_1,a_2,0,0,\dots,0)$, $\dots$, $(a_1,a_2,\dots,a_n)$. Then the probability of $\sigma$ is

$$P\{\sigma\} = \frac{\mathrm{vol}(\sigma)}{\mathrm{vol}(\Omega)} = \frac{\frac{1}{n!}\prod_{i=1}^{n} a_i}{\prod_{i=1}^{n} a_i} = \frac{1}{n!} .$$

Now, by definition $f_S(x/X_1/a_1 > X_2/a_2 > \dots > X_n/a_n)$

$$= \lim_{h\to 0} \frac{P\{(X_1 + X_2 + \dots + X_n \in (x,x + h)) \cap (X_1/a_1 > X_2/a_2 > \dots > X_n/a_n)\}}{h\, P\{X_1/a_1 > X_2/a_2 > \dots > X_n/a_n\}}$$

$$= \frac{n!}{\prod_{i=1}^{n} a_i} \lim_{h\to 0} \frac{\mathrm{vol}\ (\alpha(x,x + h) \cap \sigma)}{h} ,$$

where $\alpha(x,\ x + h) = \{(u_1,u_2,\dots,u_n) :\ u_1 + u_2 + \dots + u_n \in [x,x + h]\}$.

If the hyperplane $u_1 + u_2 + \dots + u_n = x$ is denoted by $\alpha(x)$ then

$$\lim_{h\to 0} \frac{\mathrm{vol}(\alpha(x,\ x + h) \cap \sigma)}{h} = \frac{1}{\sqrt{n}}\ \mathrm{vol}_{n-1}(\alpha(x) \cap \sigma) ,$$

where $\mathrm{vol}_{n-1}$ is the $(n - 1)$ - dimensional volume.

Thus $f_S(x/X_1/a_1 > X_2/a_2 > \dots > X_n/a_n) = \dfrac{n!}{\prod_{i=1}^{n} a_i} \dfrac{\mathrm{vol}_{n-1}(\alpha(x) \cap \sigma)}{\sqrt{n}}$.

Let us assume that $\mathrm{vol}(\Omega) = \prod_{i=1}^{n} a_i = n!$. Then

$$f_S(x/X_1/a_1 > X_2/a_2 > \dots > X_n/a_n) = \frac{1}{\sqrt{n}} \mathrm{vol}_{n-1}(\alpha(x) \cap \sigma) \tag{2.1}$$

and $\mathrm{vol}(\sigma) = 1$. Now we can apply Theorem 2 of Curry and Schoenberg (1966) which states: The fundamental spline function $M_n(x;x_0,x_1,\dots,x_n)$ is the linear density function obtained by projecting orthogonally onto the x-axis the volume of an n-dimensional simplex $\sigma_n$, of volume unity, so located that its $n+1$ vertices project orthogonally into the points $x_0,x_1,\dots,x_n$ of the x-axis, respectively.

Let us project orthogonally the simplex $\sigma$ onto the g-axis with a parametric equation $u_1 = t/\sqrt{n}$, $u_2 = t/\sqrt{n}$, ..., $u_n = t/\sqrt{n}$, $t \subset R$. If the coordinate system $0,g$ is introduced onto the g-axis so that the point $(t/\sqrt{n}, t/\sqrt{n}, \dots, t/\sqrt{n}) \in g$ has the coordinate $t$ then the orthogonal projection of the vertex $A_i(a_1,a_2\dots,a_i,0,0,\dots,0)$ of the simplex $\sigma$ onto the g-axis will be the point $\bar{A}_i$ with the coordinate $(a_1 + a_2 + \dots + a_i)/\sqrt{n}$, $i = 1,2,\dots,n$, with respect to $0,g$. Then according to the above theorem of Curry and Schoenberg

$$\mathrm{vol}_{n-1}(\alpha(x) \cap \sigma) = M_n(x;0, a_1/\sqrt{n}, (a_1 + a_2)/\sqrt{n},\dots,(a_1 + a_2\dots + a_n)/\sqrt{n}) .$$

Now from (1.1), setting $a_0 \equiv 0$ we have $\mathrm{vol}_{n-1}(\alpha(x) \cap \sigma)$

$$= n \sum_{j=0}^{n} \frac{((a_0 + a_1 + \dots + a_j)/\sqrt{n} - x/\sqrt{n})_+^{n-1}}{\prod_{\substack{i=0\\ i\neq j}}^{n} ((a_0 + a_1 + \dots + a_j)/\sqrt{n} - (a_0 + a_1 + \dots + a_i)/\sqrt{n})}$$

$$= \sqrt{n}\, M_n(x;0,a_1,a_1 + a_2, \dots, a_1 + a_2 + \dots + a_n) .$$

From (2.1) for the case $\prod_{i=1}^{n} a_i = n!$ we have

$$f_S(x/X_1/a_1 > \dots > X_n/a_n) = M_n(x;0,a_1,a_1 + a_2, \dots, a_1 + a_2 + \dots + a_n).$$

Let $\mathrm{vol}(\Omega) = \prod_{i=1}^{n} a_i = k \neq n!$, $k > 0$. By the simple transformation $\tilde{X}_i = X_i d$, $d = \sqrt{n!/k}$, $i = 1,2,\ldots,n$ we obtain the set of random variables $\tilde{X}_1, \tilde{X}_2, \ldots, \tilde{X}_n$ uniformly distributed on $(0,b_i)$, $b_i = a_i d$, $i = 1,2,\ldots,n$ and such that $\mathrm{vol}(\Omega) = \prod_{i=1}^{n} b_i = n!$.

Then

$$(2.2) \qquad f_{\tilde{S}}(x/\tilde{X}/b_1 > \tilde{X}/b_2 > \ldots > \tilde{X}/b_n) = M_n(x;0,b_1,b_1 + b_2,\ldots,b_1 + b_2 + \ldots + b_n) .$$

Changing back to the variables $X_1,X_2,\ldots,X_n$ on the left side of (2.2) and applying (1.1) to the right side we have

$$f_S(\bar{x}/X_1/a_1 > X_2/a_2 > \ldots > X_n/a_n) = M_n(\bar{x};0,a_1,a_1 + a_2,\ldots,a_1 + a_2 + \ldots + a_n),$$

where $\bar{x} = \frac{x}{d}$, proving the assertion of Lemma 2.1.

Remark: For clarity and simpler notation we have confined ourselves to the case $a_i > 0$, $i = 1,\ldots,n$. Since the divided difference (1.1) is a symmetric function of its arguments ie., it is independent on the order of the points $x_0,x_1,\ldots,x_n$ in the argument list it can be easily seen that the above Lemma is true for $a_i$ real.

We are now ready to prove

THEOREM 2.1.

Let $X_1,X_2,\ldots,X_n$ be independent random variables and $X_i$ be uniformly distributed on the interval $(0,a_i)$, $a_i$ - real, $i = 1,2,\ldots,n$. Let $S = X_1 + X_2 + \ldots + X_n$.
Then

$$(2.3) \quad f_S(x) = \frac{1}{n!} \sum_{(i_1,i_2,\ldots,i_n) \subset Q_n} M_n(x;0,a_{i_1},a_{i_1} + a_{i_2},\ldots,a_{i_1} + a_{i_2} + \ldots + a_{i_n}) ,$$

where $Q_n$ are the permutations of $(1,2,\ldots,n)$.

PROOF.

$$f_S(x) = \sum_{(i_1,i_2,\ldots,i_n) \subset Q_n} P\{X_{i_1}/a_{i_1} > X_{i_2}/a_{i_2} > \ldots > X_{i_n}/a_{i_n}\} \cdot$$

$$f_S(x/X_{i_1}/a_{i_1} > X_{i_2}/a_{i_2} > \ldots > X_{i_n}/a_{i_n})$$

$$= \sum_{(i_1,i_2,\ldots,i_n) \subset Q_n} \frac{1}{n!} f_S(x/X_{i_1}/a_{i_1} > X_{i_2}/a_{i_2} > \ldots > X_{i_n}/a_{i_n})$$

$$= \frac{1}{n!} \sum_{(i_1,i_2,\ldots,i_n) \subset Q_n} M_n(x;0,a_{i_1},a_{i_1} + a_{i_2},\ldots,a_{i_1} + a_{i_2} + \ldots + a_{i_n})$$

by Lemma 2.1.

Another interesting representation of $f_s(x)$ arises if the notion of discrete B-splines as viewed by C. de Boor (1976) is assumed.

Let $0 < x_1 < \ldots < x_n$, $a_1 = x_1,\ldots,a_n = x_n - x_{n-1}$ and $\underline{\tau}(i_1,\ldots,i_n) =$ $= \{0,a_{i_1},a_{i_1} + a_{i_2},\ldots,a_{i_1} + \ldots + a_{i_n}\}$ where $(i_1,\ldots,i_n) \subset Q_n$. We shall also use the notation $\underline{\tau}(i_1,\ldots,i_n) = \{\tau_{0,(i_1,\ldots,i_n)},\ldots,\tau_{n,(i_1,\ldots,i_n)}\}$ in which for the sake of brevity we will further drop the second subscript and write $\{\tau_0,\ldots,\tau_n\}$.

Let $\underline{t} = \{t_1,t_2,\ldots,t_p\}$, $p = \sum_{i=0}^{n} \binom{n}{i}$ be the strictly increasing sequence obtained by ordering the set

$$\{0,a_{i_1},a_{i_1} + a_{i_2},\ldots,a_{i_1} + a_{i_2} + \ldots + a_{i_j},\ldots,a_{i_1} + a_{i_2} + \ldots + a_{i_n}\},$$

where $(i_1,\ldots,i_j) \subset C_j^n$, $C_j^n$ are the combinations of $j$ from $n$, $j = 1,\ldots n$.

The B-splines of order $n$ for the knot sequence $\underline{t}$ are given by

$M_{i,n,\underline{t}}(x) = n[t_i,\ldots,t_{i+n}](\cdot - x)_+^{n-1}$, $i = 1,\ldots,q$, where $[t_i,\ldots,t_{i+n}]$ denotes a n-th order divided difference and $q = p-n$.

Since $M_n(x;0, a_{i_1},a_{i_2} + a_{i_2},\ldots,a_{i_1} + a_{i_2} + \ldots + a_{i_n})$ of (2.3) is also a spline with knots $\underline{t}$ it can be represented as a linear combination of the B-splines $M_{i,n,\underline{t}}(x)$,

$i = 1,\ldots q$ by the Curry-Schoenberg (1966) theorem (see p. 113). The non-negative coefficients of such a linear combination termed by C. de Boor (1976) discrete B-splines are defined as

$$(2.4)\qquad \beta_{j,n,\underline{\tau}(i_1,\ldots,i_n),\underline{t}}(i) = (t_{i+n}-t_i)[\tau_j,\ldots,\tau_{j+n}](\cdot - t_{i+1})_+ \cdots (\cdot - t_{i+n-1})_+ .$$

Remark 1: When $j = 0$ we simply write $\beta_{n,\underline{\tau}(i_1,\ldots,i_n),\underline{t}}$.

Remark 2: When $n = 1$, (2.4) reads

$$\beta_{j,1,\underline{\tau}(i_1,\ldots,i_n),\underline{t}}(i) = (t_{i+1}-t_i)[\tau_j,\tau_{j+1}](\cdot - t_i)_+^0$$

where

$$(\tau_r - t_i)_+^0 = \begin{cases} 1 , & \tau_r > t_i \\ 0 , & \tau_r < t_i \end{cases} .$$

Thus the density $f_S(x)$ is expressed in terms of discrete B-splines as

$$(2.5)\qquad f_S(x) = \frac{1}{n!} \sum_{i=1}^{q} \sum_{(i_1,\ldots,i_n)\subset Q_n} \beta_{n,\underline{\tau}(i_1,\ldots,i_n),\underline{t}}(i) M_{i,n,\underline{t}}(x) .$$

Note that if $t_j \le x \le t_{j+1}$, $j = n,\ldots,q-1$ then

$$f_S(x) = \frac{1}{n!} \sum_{i=j-n+1}^{j} \sum_{(i_1,\ldots,i_n)\subset Q_n} \beta_{n,\underline{\tau}(i_2,\ldots,i_n),\underline{t}}(i) M_{i,n,\underline{t}}(x) ,$$

since other $M_{i,n,\underline{t}}(x)$'s vanish identically. This means that the evaluation of $f_S(x)$ for any $x \in [x_0,x_n]$ requires summation of no more than $n$ of the $q$ terms on the righthand side of (2.5).

The discrete B-splines in (2.5) can be determined from the recurrence expression

$$\beta_{j,n,\underline{\tau}(i_1,\ldots,i_n),\underline{t}}(i) = \alpha_{j,i,n}\big[(\tau_{j+n} - t_{i+n-1})\beta_{j+1,n-1}(i)$$

$$(2.6)\qquad + (t_{i+n-1} - \tau_j)\beta_{j,n-1}(i)\big] , \text{ where}$$

$$\alpha_{j,i,n} = (t_{i+n} - t_i)/[(\tau_{j+n} - \tau_j)(t_{i+n-1} - t_i)] .$$

Remark: For a proof of (2.6) with differently normalized discrete B-splines see Jia (1983).

## 3. B-splines and order statistics.

We first state the following Lemma, providing a probabilistic interpretation of the B-spline in terms of order statistics. It should be noted that this result easily follows differentiating the distribution formula given by Dempster and Kleyle (1968) and Ali (1969). For the sake of consistency with the rest of this note another proof is given below.

LEMMA 3.1.

Let $Z_1, Z_2, \ldots, Z_n$ be $n$ independent uniformly distributed on $(0,1)$ random variables and let $Z_{(1)} > Z_{(2)} > \ldots > Z_{(n)}$ be their order statistics. Let $a_1, a_2, \ldots, a_n$ be fixed real numbers. Then the density of the linear combination $L = a_1 Z_{(1)} + a_2 Z_{(n)} + \ldots + a_n Z_{(n)}$ is $f_L(x) = M_n(x; 0, a_1, a_1+a_2, \ldots, a_1+a_2+\ldots+a_n)$.

PROOF.

From Lemma 2.1 we have

$$f_{a_1Z_1+a_2Z_2+\ldots+a_nZ_n}(x \,/\, Z_1 > Z_2 > \ldots > Z_n)$$
$$= M_n(x; 0, a_1, a_1+a_2, \ldots, a_1+a_2+\ldots+a_n) \quad .$$

It can be easily seen that

$$f_{a_1Z_1+a_2Z_2+\ldots+a_nZ_n}(x \,/\, Z_1 > Z_2 > \ldots > Z_n)$$
$$= \lim_{h \to 0} \frac{P\{(a_1Z_1+a_2Z_2+\ldots+a_nZ_n \in (x,x+h)) \cap (Z_1 > Z_2 > \ldots > Z_n)\}}{h\, P\{Z_1 > Z_2 > \ldots > Z_n\}}$$

$$= \lim_{h \to 0} \frac{n!}{h} \int \underset{\Delta}{\ldots} \int 1 \, du_1 \ldots du_n = \lim_{h \to 0} \frac{1}{h} \int \underset{\Delta_1 \cap \Delta_2}{\ldots} \int n! \, du_1 \ldots du_n \quad , \tag{3.1}$$

where

$$\Delta = \{(u_1, u_2, \ldots, u_n) \;:\; 1 > u_1 > u_2 > \ldots > u_n > 0 \;,\; x < a_1u_1 + a_2u_2 + \ldots + a_nu_n < x+h\} \quad ,$$

$$\Delta_1 = \{(u_1, u_2, \ldots, u_n) \;:\; 1 > u_1 > \ldots > u_n > 0\} \quad ,$$

$$\Delta_2 = \{(u_1, u_2, \ldots, u_n) \;:\; x < a_1u_1 + a_2u_2 + \ldots + a_nu_n < x+h\} \quad .$$

Denote

$$\zeta_{\Delta_1} = \begin{cases} 1 & , \quad (u_1,u_2,\ldots,u_n) \in \Delta_1 \\ 0 & , \quad (u_1,u_2,\ldots,u_n) \bar{\in} \Delta_1 \end{cases} .$$

Then (3.1) yields

(3.2) $$= \lim_{h \to 0} \frac{1}{h} \int_{\Delta_2} \cdots \int \zeta_{\Delta_1} \, n! \, du_1 \ldots du_n .$$

Since the density of the random vector $(Z_{(1)},Z_{(2)},\ldots,Z_{(n)})$ is $\zeta_{\Delta_1} n!$ (c.f. S. Karlin and H. Taylor, (1981), p. 102)

(3.3) $$\frac{1}{h} \int_{\Delta_2} \cdots \int \zeta_{\Delta_1} \, n! \, du_1 \ldots du_n = \frac{P\{a_1 Z_{(1)} + a_2 Z_{(2)} + \ldots + a_n Z_{(n)} \in (x,x+h)\}}{h} .$$

From (3.3) replacing in (3.2) we finally obtain

(3.4) $$\lim_{h \to 0} \frac{1}{h} \int_{\Delta_2} \cdots \int \zeta_{\Delta_1} \, n! \, du_1 du_2 \ldots du_n = \lim_{h \to 0} \frac{P\{a_1 Z_{(1)} + a_2 Z_{(2)} + \ldots + a_n Z_{(n)} \in (x,x+h)\}}{h} = f_L(x) .$$

Since the vector $(Z_{(1)},Z_{(2)},\ldots,Z_{(n)})$ has a density, the random variable $a_1 Z_{(1)} + a_2 Z_{(2)} + \ldots + a_n Z_{(n)}$ also has a density, denoted by $f_L(x)$ in (3.4). That completes the proof of Lemma 3.1.

Remark 1: It can be easily seen by a linear transformation that if $x_0 < x_1 < \ldots < x_n$, $a_1 = x_1 - x_0$, $a_2 = x_2 - x_1, \ldots, a_n = x_n - x_{n-1}$ and $\tilde{Z}_1, \tilde{Z}_2, \ldots, \tilde{Z}_n$ are independent random variables uniformly distributed in the interval $(x_0/(x_n - x_0) \, , \, x_0/(x_n - x_0) + 1)$ then the linear combination $\tilde{L} = a_1 \tilde{Z}_{(1)} + a_2 \tilde{Z}_{(2)} + \ldots + a_n \tilde{Z}_{(n)}$ has the density $f_{\tilde{L}}(x) = M_n(x; x_0, x_1, \ldots, x_n)$. Since $\tilde{Z}_i = Z_i + x_0/(x_n - x_0)$, $i = 1,2,\ldots,n$ and $\tilde{Z}_{(i)} = Z_{(i)} + x_0/(x_n - x_0)$, $i = 1,2,\ldots,n$, $f_{\tilde{L}}(x) = f_{L+x_0}(x) = M_n(x; x_0, x_1, \ldots, x_n)$ holds.

Remark 2. If some $a_i$, $i = 1,\ldots,n$ are zero i.e. the B-spline has multiple knots than the above Lemma still holds. The righthand side of formula (1.1) must be replaced by the expression for divided differences for multiple arguments (c.f. Shumaker (1981) p. 119). For stable computation of $f_L(x)$ however the recurrence relations given next should be

Olds, E. G. (1952). A note on the convolution of uniform distributions. Ann. Math. Statist. 23 282-285.

Shumaker, L. L. (1981). Spline functions: Basic Theory. Wiley, New York.

van Zwet, W. R. (1979). The Edgeworth Expansion for linear combinations of uniform order statistics, in: Proceedings of the Second Prague Symposium on Asymptotic Statistics, eds. P. Mandl, and M. Huskova, Prague: Charles University, 93-101.

Weisberg, H. (1971). The distribution of linear combinations of order statistics from the uniform distribution. Ann. Math. Statist. 42, 704-9.

ZGI/VKK/jvs

REFERENCES

Ali, MiZ M. (1969). Distribution of linear combination of order statistics from rectangular population. Bull. Inst. Statist. Res. Tr. 3, 1 1-21.

de Boor, C. (1972). On calculating with B-splines. J. Approximation Theory. 6 50-62.

de Boor, C. (1976). Splines as linear combinations of B-splines, in: Approximation Theory II, ed. by G. G. Lorentz, C. K. Chui and L. L. Schumaker, Academic Press, New York, 1-47.

de Boor, C. and De Vore, R. (1984). A geometric proof of total positivity for spline interpolation. University of Wisconsin, Mathematics Research Center Tech. Summary Report #2677.

Cicchitelli, G. (1976). The Sampling Distribution of Linear Combinations of Order Statistics. Metron, 34, 269-302.

Cox, M. G. (1972). The numerical evaluation of B-splines. J. Inst. of Math. and Appl. 10, 2 139-149.

Currie, I. D. (1981). On the distributions determined by random variables distributed over the n-cube. Ann. Statist. 9, 822-833.

Curry, H. B. and Schoenberg, I. J. (1966). On Polya frequency functions. IV The fundamental spline functions and their limits, J. d'Analyse Math. 17 71-107.

David, H. A. (1981). Order Statistics. Wiley, New York.

Dempster, A. P. and Kleyle, R. M. (1968). Distributions determined by cutting a simplex with hyperplanes. Ann. Math. Statist. 39, 5 147?-8.

Hecker, H. (1976). A characterization of the asymptotic normality of linear combinations of order statistics from the uniform distribution. Ann. Statist. 4, 1244-1246.

Jia, Rong-Quing (1983). Total Positivity of the Discrete Spline Collocation Matrix. J. Approx. Theory, 39, 11-23.

Karlin, S. and Taylor, H. M. (1981). A second course in stochastic processes. Academic Press. New York.

Neuman, E. (1980). Moments of B-splines, in: Approx. Theory III, p. 675-8, Academic Press, New York.

Application.

The circular serial correlation coefficient with lag A is given by

$$(5.1) \qquad {}_AR_N = \sum_{j=1}^{N} (X_j-\bar{X})(X_{j+A}-\bar{X}) / \sum_{j=1}^{N} (X_j-\bar{X})^2 ,$$

where $X_{N+j} = X_j$ and $A <$ sample size $N$.

It may be noted that the marginal density of ${}_AR_N$, in the case $X_j$, $1,\dots,N$ i.i.d. $N(0,1)$ and N-odd is given by the B-spline $M_n(x; {}_AC_1, {}_AC_2, \dots, {}_AC_{n+1})$, where ${}_AC_1 > {}_AC_2 > \dots > {}_AC_{n+1}$ are the distinct latent roots of the matrix of the quadratic form in the numerator of (5.1), $n+1 = (N-1)/2$. This follows differentiating and slightly rewriting formula (1.1) of Dempster and Kleyle (1968) which gives the marginal distribution of ${}_AR_N$.

Let $T_{2n} = \sqrt{3/n}\,(Z_{1,2n} + \ldots + Z_{2n,2n}) - \sqrt{3n}$. Then, for the density of $T_{2n}$ we have

$$\frac{d}{dx} P\left(\sqrt{2n}\,\frac{X_1 + \ldots + X_{2n}}{2n} < x\right) = f_{T_{2n}}(x)$$

$$= M_{2n}(x; -n\sqrt{3/n},\ (-n+1)\sqrt{3/n},\ (-n+2)\sqrt{3/n}\ ,\ldots,\ n\sqrt{3/n})$$

Then according to the LCLT

$$\lim_{n\to\infty}\ \sup_{-\infty<x<\infty} \left| M_{2n}(x; -n\sqrt{3/n},\ (-n+1)\sqrt{3/n},\ (-n+2)\sqrt{3/n}\ ,\ldots,\ n\sqrt{3/n}) - \frac{1}{\sqrt{\pi}}\exp(-x^2) \right| = 0 \quad .$$

**EXAMPLE 5.**

$$T_{2k-1} = -\frac{2k}{1} + 2k\,\frac{2}{1\cdot 3}\,Z_{1,2k-1} + 2k\,\frac{2}{3\cdot 5}\,Z_{2,2k-1} + 2k\,\frac{2}{5\cdot 7}\,Z_{3,2k-1}$$

$$+\ldots+ 2k\,\frac{2}{(2k-3)(2k-1)}\,Z_{k-1,2k-1} + 2k\,\frac{2}{(2k-1)}\,Z_{k,2k-1} +$$

$$+ 2k\,\frac{2}{(2k-3)(2k-1)}\,Z_{k+1,2k-1} + 2k\,\frac{2}{(2k-5)(2k-3)}\,Z_{k+2,2k-1} +\ldots+ 2k\,\frac{2}{1\cdot 3}\,Z_{2k-1,2k-1} \quad .$$

Then, by Lemma 3.1

$$f_{T_{2k-1}}(x) = M_n(x; -\frac{2k}{1}\ ,\ -\frac{2k}{3}\ ,\ldots,\ -\frac{2k}{2k-1}\ ,\ \frac{2k}{2k-1}\ ,\ \frac{2k}{2k-3}\ ,\ldots,\ \frac{2k}{1})$$

and by Example 5 of Curry and Schoenberg (1966)

$$\lim_{k\to\infty}\ \sup_{-\infty<x<\infty} \left| f_{T_{2k-1}}(x) - \frac{1}{\pi\ \cosh\ x} \right| = 0 \quad .$$

**EXAMPLE 6.**

Let $T_{2k-1} = -2\log k + (\frac{2k}{k} + 2\log k)\,Z_{k,2k-1} + (\frac{2k}{k-1} - \frac{2k}{k})Z_{k+1,2k-1} +$ $+ (\frac{2k}{k-2} - \frac{2k}{k-1})Z_{k+2,2k-1} +\ldots+ (\frac{2k}{1} - \frac{2k}{2})Z_{2k-1,2k-1} = -2\log k + 2(1+\log k)Z_{k,2k-1} +$ $+ 2k\,\frac{1}{k(k-1)}\,Z_{k+1,2k-1} + 2k\,\frac{1}{(k-1)(k-2)}\,Z_{k+2,2k-1} +\ldots+ 2k\,\frac{1}{2\cdot 1}\,Z_{2k-1,2k-1}$. Then by Lemma 3.1 the density

$$f_{T_{2k-1}}(x) = M_{2k-1}(x; \overbrace{-2\log k,\ldots,-2\log k}^{k},\ \frac{2k}{k}\ ,\ \frac{2k}{k-1}\ ,\ldots,\ \frac{2k}{1}) \quad \text{and}$$

$$\lim_{k\to\infty}\ \sup_{-\infty<x<\infty} \left| f_{T_{2k-1}}(x) - \exp(-\exp(-x) - x) \right| = 0 \quad ,$$

by Example 6 of Curry and Schoenberg (1966).

$$f_{T_n}(x) \text{ converges uniformly to } \lambda(x) = \begin{cases} \exp(-x) & , \quad x > 0 \\ 0 & , \quad x < 0 \end{cases}$$

for $x$ outside an arbitrarily small neighborhood of the point $x = 0$.

**EXAMPLE 3.**

Let $T_{2k-1} = -\sqrt{k} + 2\sqrt{k}\, Z_{k,2k-1}$. Then, by Lemma 3.1

$$f_{T_{2k-1}}(x) = M_{2k-1}(x; \overbrace{-\sqrt{k},\ldots,-\sqrt{k}}^{k}, \overbrace{\sqrt{k},\ldots,\sqrt{k}}^{k})$$

converges uniformly for all real $x$ to $\exp(-x^2)/\sqrt{\pi}$.

More generally, by a theorem of Mosteller (c.f. David (1981) p. 255) it is known that if $0 < p < 1$ and

$$T_n = -\sqrt{\frac{np}{1-p}} + \frac{n}{\sqrt{np(1-p)}} Z_{n-[np],n}$$

then $T_n$ converges weakly to the random variable $T$ normally distributed with $ET = 0$ and $DT = 1$. Hence, by Lemma 3.1

$$\int_{-\infty}^{x} M_n(u; \overbrace{-\sqrt{\frac{np}{1-p}},\ldots,-\sqrt{\frac{np}{1-p}}}^{n-[np]}, \overbrace{\frac{n(1-p)}{\sqrt{np(1-p)}},\ldots,\frac{n(1-p)}{\sqrt{np(1-p)}}}^{[np]+1})du \xrightarrow[n\to\infty]{} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} \exp(-u^2/2)du$$

for every $x \in R_1$.

**EXAMPLE 4.**

According to the local central limit theorem (LCLT) if $X_n$, $n = 1,2,\ldots$ are independent with a common distribution $F$ having mean $0$ and variance $\sigma^2$ and if $F$ has a bounded density, then

$$\lim_{n\to\infty} \sup_{-\infty<x<\infty} \left| \frac{d}{dx} P(\sqrt{n}\,\bar{X}_n < x) - \frac{1}{\sqrt{2\pi}} \exp(-x^2/2\sigma^2) \right| = 0 \quad .$$

Let $X_n = \sqrt{6}\,(Z_n - 1/2)$. Then

$$F(x) = \begin{cases} 1 & , \quad x > 1/2\sqrt{6} \\ x/\sqrt{6} + 1/2 & , \quad -1/2\sqrt{6} < x < 1/2\sqrt{6} \\ 0 & , \quad x < -1/2\sqrt{6} \end{cases}$$

and the conditions of the LCLT hold with $\sigma^2 = 1/2$. Thus we have

$$\sqrt{2n}\,\frac{X_1+\ldots+X_{2n}}{2n} = \sqrt{3/n}\,(Z_{1,2n}+\ldots+Z_{2n,2n}) - \sqrt{3n} \quad .$$

## 5. Examples and applications.

The following examples illustrate how, on the basis of Lemma 3.1, providing a probabilistic interpretation of the B-spline, asymptotic results for B-splines (Curry and Schoenberg (1966), Section 9) may be applied to linear combinations of order statistics. It is also shown (examples 3, 4) that using limit theorems of probability theory the results of Curry and Schoenberg (1966) (examples 3, 4) can be derived, even in slightly more generality (example 3).

As above, here $Z_{1,n} > Z_{2,n} > \ldots > Z_{n,n}$ is the order statistics from the uniform distribution on $(0,1)$.

EXAMPLE 1.

Let $x_0 < x_1 < \ldots$ be a convergent sequence of reals with $\lim_{n\to\infty} x_n = B$. Then the sequence of random variables $T_n = x_0 + (x_1-x_0)Z_{1,n} + (x_2-x_1)Z_{2,n} + \ldots + (x_n-x_{n-1})Z_{n,n}$ converges weakly to the constant $B$.

PROOF.

According to Lemma 3.1 and Lemma 6 of Curry and Schoenberg (1966) for the density $f_{T_n}(x)$ of the random variable $T_n$ we have

$$\int_{-\infty}^{\infty} f_{T_n}(x)(1 + xs/(n+1))^{-n-1}dx = \int_{-\infty}^{\infty} M_n(x)(1 + xs/(n+1))^{-n-1}dx$$

$$= \frac{1}{\prod_{i=0}^{n} (1 + x_i s/(n+1))} .$$

It can be easily seen that

$$\lim_{n\to\infty} \frac{1}{\prod_{i=0}^{n}(1 + x_i s/(n+1))} = e^{-Bs} .$$

Hence, by Lemma 4 of Curry and Schoenberg (1966) $T_n$ converges weakly to $B$ and therefore $T_n$ converges to $B$ in probability.

EXAMPLE 2.

Let $T_n = nZ_{n,n}$. Then, by Lemma 3.1 the density of the random variable $T_n$ $f_{T_n}(x) = M_n(x;0,0,\ldots,0,n)$ and hence, by Theorem 7 of Curry and Schoenberg

The remainder term in the last expansion is given by van Zwet (1979) to be typically of order $n^{-3/2}$.

We finally remark that the Berry-Essenn bound of Corollary 1 can be sharpened by means of Corollary 2 to obtain

$$\sup_x \left| \int_{-\infty}^{x} M_n(t; y_{0,n}, y_{1,n}, \dots, y_{n,n})dt - \right.$$

$$\left. - \Phi(x) \right| \le C\left[\left| \sum_{j=0}^{n} \{y_{j,n}/\sqrt{(n+1)(n+2)}\}^3 \right| + \sum_{j=0}^{n} \{y_{j,n}/\sqrt{(n+1)(n+2)}\}^4\right] .$$

Following van Zwet (1979) we develop the expansion of Theorem 4.4 for the special cases $m = 3$ and $m = 5$. For $m = 3$ it reduces to a bound of Berry-Essenn type, typically of order $n^{-1/2}$.

COROLLARY 1.

There exists a constant $C$ such that for every $n = 1,2,\ldots$ and every $a_1(n),\ldots,a_n(n)$ with $\sum_{j=1}^{n} a_j^2(n) \neq 0$,

$$\sup_x \left| \int_{-\infty}^{x} M_n(t; y_{0,n}, y_{1,n}, \ldots, y_{n,n})dt - \Phi(x) \right| \leq C \sum_{j=0}^{n} |y_{j,n}/\sqrt{(n+1)(n+2)}|^3 .$$

Let now $m = 5$. Define

$$F_n(x) = \Phi(x) - \phi(x)\Big[\frac{1}{3} \sum_{j=0}^{n} [y_{j,n}^3/\sqrt{(n+1)(n+2)}]^3 \cdot H_2(x) +$$

$$+ \Big\{\frac{1}{4} \sum_{j=0}^{n} [y_{j,n}/\sqrt{(n+1)(n+2)}]^4 - \frac{1}{2n}\Big\} H_3(x) +$$

$$+ \frac{1}{18} \Big\{ \sum_{j=0}^{n} [y_{j,n}/\sqrt{(n+1)(n+2)}]^3 \Big\}^2 H_5(x) \Big] .$$

COROLLARY 2.

There exists a constant $C$ such that for every $n = 1,2,\ldots$ and every $a_1(n),\ldots,a_n(n)$ with $\sum_{j=1}^{n} a_j^2(n) \neq 0$,

$$\sup_x \left| \int_{-\infty}^{x} M_n(t; y_{0,n}, y_{1,n}, \ldots, y_{n,n})dt - \right.$$

$$\left. - F_n(x) \right| \leq C \sum_{j=0}^{n} |y_{j,n}/\sqrt{(n+1)(n+2)}|^5 .$$

Corollaries 1 and 2 follow by Lemma 3.1 and Corolaries 1 and 2 of van Zwet (1979).

$$V_{j,n}(x) = x_{j,n} - \bar{x}_n - \sigma_n x \quad ,$$

$$W_{j,n}(x) = V_{j,n}(x)\{\sum_{j=0}^{n} V_{j,n}^2(x)\}^{-1/2}$$

$$= \left\{\frac{n+2}{n+2+x^2}\right\}^{1/2} (y_{j,n} - x)\{(n+1)(n+2)\}^{-1/2} \quad ,$$

$$\xi(x) = -\sum_{j=0}^{n} W_{j,n}(x) = x\left\{\frac{(n+1)}{n+2+x^2}\right\}^{1/2} \quad .$$

It is easy to check that $\Sigma W_{j,n}^2(x) = 1$. For integer $m \geq 3$ and real $x$ and $z$, let

$$G_{m,n}(z,x) = \Phi(x) - \phi(z)\sum^{*} H_{(\Sigma k\nu_k - 1)}(z) \cdot \prod_{k=3}^{m-1} \frac{1}{\nu_k!} \{\frac{1}{k} \sum_{j=0}^{n} W_{j,n}^k(x)\}^{\nu_k} \quad ,$$

where $\sum^{*}$ denotes summation over all nonnegative integers $\nu_3,\ldots,\nu_{m-1}$ with

$$1 \leq \sum_{k=3}^{m-1} (k-2)\nu_k \leq m-3 \quad .$$

**THEOREM 4.4.**

For every integer $m \geq 3$ there exists a constant $C_m$ such that for every $n = 1,2,\ldots$ and every $a_1(n),\ldots,a_n(n)$ with $\Sigma a_j^2(n) \neq 0$,

$$\sup_x \left| \int_{-\infty}^{x} M_n(t;y_{0,n},y_{1,n},\ldots,y_{n,n})dt - \right.$$

$$\left. - G_{m,n}(\xi(x),x) \right| \leq C_m \sum_{j=0}^{n} |y_{j,n}/\sqrt{(n+1)(n+2)}|^m \quad .$$

**PROOF.**

Follows by Lemma 3.1 and the theorem of van Zwet (1979) establishing the Edgeworth expansion for the distribution function of $T_n$.

Since we require the condition (4.4) to be fulfilled the remainder in the above expansion tends to zero for every $m \geq 3$ and for increasing $m$. The rates will depend of course on the triangular array $\{a_j(n)\}$ related to the knots of the B-spline. The remainder approximately behave as $n^{-1/2(m-2)}$ (see van Zwet (1979)).

$$y_{j,n} = (x_{j,n} - \bar{x}_n)/\sigma_n, \; j = 0,\ldots,n \quad .$$

Since $\Sigma\, a_j^2(n) \neq 0$ implies that $\sigma_n^2 \neq 0$ the $y_{j,n}$'s are well defined with $\Sigma\, y_{j,n} = 0$ and $\Sigma\, y_{j,n}^2 = (n+1)(n+2)$.

**THEOREM 4.3.**

$$\lim_{n\to\infty} \sup_{-\infty<x<\infty} \left| M_n(x; y_{0,n}, y_{1,n}, \ldots, y_{n,n}) - \frac{1}{\sqrt{2\pi}} \exp(-x^2/2) \right| = 0$$

if and only if

$$\lim_{n\to\infty} \left[ \max_{0\leq j\leq n} \left| y_{j,n} / \sqrt{(n+1)(n+2)} \right| \right] = 0 \quad . \tag{4.4}$$

**PROOF.**

Applying Lemma 3.1 to the theorem of Hecker (1976) which gives a necessary and sufficient condition for the asymptotic normality of the linear combination $T_n = y_{0,n} + (y_{1,n}-y_{0,n})Z_{1,n} + (y_{2,n}-y_{1,n})Z_{2,n} + \ldots + (y_{n,n}-y_{n-1,n})Z_{n,n}$ we easily obtain the assertion of Theorem 4.3.

Under the condition (4.4) we shall derive an expansion for the spline distribution function

$$\int_{-\infty}^{x} M_n(t; y_{0,n}, y_{1,n}, \ldots, y_{n,n})dt$$

using a theorem of van Zwet (1979).

Define

$$\Phi(x) = \int_{-\infty}^{x} \phi(t)dt \quad ,$$

$\phi(x) = \dfrac{1}{\sqrt{2\pi}} \exp(-x^2/2)$,

$H_r$ - the Hermite polynomial of degree $r$ i.e. $\dfrac{d^r\phi(x)}{dx^r} = (-1)^r \phi(x) H_r(x)$. For $j = 0,\ldots,n$ and real $x$ let

THEOREM 4.2.

Let the sequence of random variables $T_n = a_1(n)Z_{(1)} + a_2(n)Z_{(2)} + \dots +$ $a_n(n)Z_{(n)} + a_{n+1}(n)$ converge in distribution to a random variable with a distribution function $F(x)$, and let

$$\int_{-\infty}^{\infty} e^{-sx}\, dF(x) \neq e^{\delta s} .$$

Hence $dF(x) = \Lambda(x)dx$, where $\Lambda(x)$ is a Polya frequency function and

$$\lim_{n\to\infty} f_{T_n}(x) = \Lambda(x) \tag{4.3}$$

uniformly for all $x$ provided that $\Lambda(x)$ is not of the form (4.2). Here $f_{T_n}(x)$ denotes the density function of the random variable $T_n$.

For $\Lambda(x) = \Lambda_0(x)$, again (4.3) holds uniformly for $x$ outside an arbitrarily small neighborhood of the point $x = \delta - \delta_1$.

If $\Phi(s) = \exp(\delta s)$ as in case i) then

$$\lim_{n\to\infty} f_{T_n}(x) = 0$$

uniformly in $x$ outside an arbitrarily small neighborhood of the point $x = \delta$.

The proofs of Theorems 4.1 and 4.2 directly follow from theorems 6 and 7 of Curry and Schoenberg (1966) respectively, applying Lemma 3.1.

We shall now use two theorems of asymptotic statistics to establish easily results for B-splines.

Let $a_1(n),\dots,a_n(n)$ be a real sequence with $x_{0,n} = 0$, $x_{1,n} = a_1(n)$, $x_{2,n} = a_1(n) + a_2(n),\dots,x_{n,n} = a_1(n) + \dots + a_n(n)$ and $\sum_{j=1}^{n} a_j^2(n) \neq 0$ for every $n$.

Define

$$\bar{x}_n = \frac{1}{n+1} \sum_{j=0}^{n} x_{j,n} ,$$

$$\sigma_n^2 = \frac{\sum_{j=0}^{n} (x_{j,n} - \bar{x}_n)^2}{(n+1)(n+2)} ,$$

## 4. Limit theorems.

Let us first state two limit theorems concerning linear combinations of order statistics. For the purpose the class of Polya distribution functions $F(x)$ will be required. These are the distribution functions having a bilateral Laplace transform of the form

$$\int_{-\infty}^{\infty} e^{-sx}\, dF(x) = 1/\Phi(s) \quad , \tag{4.1}$$

where $\Phi(s) = \exp(-\gamma s^2 + \delta s) \prod_{i=1}^{\infty} (1+\delta_i s)\exp(-\delta_i s)$ , $\gamma \geq 0$, $\delta, \delta_i$ real , $\Sigma\, \delta_i^2 < \infty$ (see Curry and Schoenberg (1966), p. 91).

THEOREM 4.1.

If $\{a_1(n), a_2(n), \ldots, a_n(n), a_{n+1}(n)\}$ is a sequence of constants such that the sequence of linear combinations of order statistics $a_1(n)Z_{(1)}+a_2(n)Z_{(2)}+\ldots+a_n(n)Z_{(n)}+a_{n+1}(n)$ converges weakly to a random variable $W$ with a distribution function $F(x)$, then $F(x)$ is a Polya distribution function. Conversely if $W$ is a random variable with a Polya distribution function $F(x)$ then there exists an appropriate sequence of constants $a_1(n), a_2(n), \ldots, a_n(n), a_{n+1}(n)$ such that $a_1(n)Z_{(1)}+a_2(n)Z_{(2)}+\ldots+a_n(n)Z_{(n)}+a_{n+1}(n)$ converges weakly to $W$.

It is known that $\exp(\gamma s^2)$ is the bilateral Laplace transform of the random variable $W$ normally distributed with $EW = 0$ and $D(W) = 2\gamma$. Consequently $\Phi(s)^{-1}$ (c.f. (4.1)) is the bilateral Laplace transform of the distribution of the sum of independent random variables $U = \delta + \sqrt{2\gamma}W + \delta_1 V_1 - \delta_1 + \delta_2 V_2 - \delta_2 + \ldots$, where $V_i$ is exponentially distributed with $EV_i = 1$, $i = 1,2,\ldots$ , $W \sim N(0,1)$, and $\delta$ real constant.

i) If $0 = \gamma = \delta_1 = \delta_2 = \ldots$ then the bilateral Laplace transform of the random variable $U$ is $\Phi(s) = \exp(\delta s)$ and no density function exists.

ii) If $\gamma = 0$ and only one of the constants $\delta_i$ is nonzero, $i = 1,2,\ldots$ then the random variable $U$ has a density function

$$\Lambda_0(x) = \begin{cases} \frac{1}{\delta_i} \exp(-(x-\delta+\delta_i)/\delta_i) & , \quad (x-\delta+\delta_i)/\delta_i > 0 \\ 0 & , \quad \text{elsewhere} \end{cases} , \tag{4.2}$$

which is discontinuous at the point $x = \delta-\delta_i$.

A generalization of (3.5) is straightforward. Let $\underline{t} = \{t_0,\ldots,t_p\}$ be a non-decreasing real sequence with $t_i < t_{i+n}$, $i = 0,\ldots,p-n$, $\{\rho_0,\ldots,\rho_n\}$ be an increasing integer sequence, and $\underline{\tau} = \{\tau_j = t_{\rho_j}\}$, $j = 0,\ldots,n$, $\rho_0 = 0$, $\rho_n = p$ be a subsequence of $\underline{t}$. Let also $a_1 = \tau_1-\tau_0$, $a_2 = \tau_2 - \tau_1,\ldots,a_n = \tau_n - \tau_{n-1}$ and $b_1 = t_1 - t_0,\ldots,b_p = t_p - t_{p-1}$ i.e. $a_1 = b_1+\ldots+b_{\rho_1}$, $a_2 = b_{\rho_1+1}+\ldots+b_{\rho_2},\ldots,$ $a_n = b_{\rho_{n-1}+1}+\ldots+b_{\rho_n}$. Then by the same argument as in deriving (2.5)

$$f_{a_1Z_{1,n}+\ldots+a_nZ_{n,n}}(x-\tau_0) =$$

$$= \sum_{i=1}^{q} \beta_{n,\underline{\tau},\underline{t}}(i-1) f_{b_1+\ldots+b_{i-1}+b_iZ_{1,n}+\ldots+b_{i+n-1}Z_{n,n}}(x-\tau_0)$$

where $q = p - n+1$, $\beta_{n,\underline{\tau},\underline{t}}(i-1)$ is the discrete B-spline given by (2.4).

Denote by $\mu_\ell(a_1Z_{1,n}+\ldots+a_nZ_{n,n})$ the $\ell$-th moment of the linear combination $L = a_1Z_{1,n}+\ldots+a_nZ_{n,n}$, $\ell = 0,1,\ldots$ and assume the sequence $\underline{x}$ is strictly increasing. Then the following recurrence formulae hold.

$$\mu_\ell(a_1Z_{1,n}+\ldots+a_nZ_{n,n})$$

(3.6)
$$= \binom{n+\ell}{n}^{-1} \sum_{j=0}^{\ell} x_n^{\ell-j}\binom{n+j-1}{j}\mu_j(a_1Z_{1,n}+\ldots+a_{n-1}Z_{n-1,n}) \quad ,$$

where the initial values are equal to

$$\mu_j(a_1,Z_{1,1}) = (x_1^{j+1} - x_0^{j+1})/[(j+1)(x_1-x_0)], \ (j = 0,1,\ldots) \quad .$$

(3.7)
$$\sum_{j=1}^{p} \binom{n}{j} \prod_{m=0}^{j-1} (x_p-x_m)\mu_{n-j}(a_1Z_{1,n}+\ldots+a_jZ_{j,n}) = x_p^n-x_0^n$$
$$, \quad (p = 0,1,\ldots,p < n) \quad .$$

(3.8)
$$\sum_{j=1}^{n} \binom{n}{j} \prod_{\substack{m=0\\ m\neq p}}^{j-1} (x_p-x_m)\mu_{n-j}(a_1Z_{1,n}+\ldots+a_jZ_{j,n}) = nx_p^{n-1}$$
$$, \quad (p = 0,1,\ldots,n-1) \quad .$$

The proof of formulae (3.6), (3.7) and (3.8) follows applying Lemma 3.1 to the corresponding formulae for the moments of B-splines (see Newman 1980).

used.

To emphasize the sample size $n$ in the sequel, we shall also use the notation $Z_{1,n} \geq Z_{2,n} \geq \ldots \geq Z_{n,n}$ for the order statistics of the sample $Z_1, Z_2, \ldots, Z_n$. Let $\underline{x} = \{x_0, x_1, \ldots, x_n\}$ be a non-decreasing real sequence with $x_0 < x_n$ and $a_1 = x_1 - x_0$, $a_2 = x_2 - x_1, \ldots, a_n = x_n - x_{n-1}$.

**LEMMA 3.2.**

Let $n \geq 2$,

$$f_{a_1 Z_{1,n} + \ldots + a_n Z_{n,n}}(x - x_0) = \frac{n}{n-1}\left[\frac{x - x_0}{x_n - x_0} f_{a_1 Z_{1,n-1} + \ldots + a_{n-1} Z_{n-1,n-1}}(x - x_0) + \right.$$

$$\left. + \frac{x_n - x}{x_n - x_0} f_{a_1 + a_2 Z_{1,n-1} + \ldots + a_n Z_{n-1,n-1}}(x - x_0)\right] .$$

**PROOF.**

The proof follows immediately applying Lemma 3.1 to the recurrence relation for stable evaluation of B-splines (de Boor (1972), Cox (1972)).

Let $\underline{t} = \{t_0, t_1, \ldots, t_{n+1}\}$ be a sequence obtained from $\underline{x}$ by addition of the point $t_\nu$ satisfying $x_{\nu-1} \leq t_\nu \leq x_\nu$, $0 \leq \nu \leq n$, $n > 0$. Let $a_1 = t_1 - t_0$, $a_2 = t_2 - t_1, \ldots, a_\nu = t_\nu - t_{\nu-1}, \ldots, a_{n+1} = t_{n+1} - t_n$.

**LEMMA 3.3.**

$$f_{a_1 Z_{1,n} + \ldots + (a_\nu + a_{\nu+1}) Z_{\nu,n} + a_{\nu+2} Z_{\nu+1,n} + \ldots + a_{n+1} Z_{n,n}}(x - t_0) =$$

$$= \frac{t_\nu - t_0}{t_{n+1} - t_0} f_{a_1 Z_{1,n} + \ldots + a_\nu Z_{\nu,n} + a_{\nu+1} Z_{\nu+1,n} + \ldots + a_n Z_{n,n}}(x - t_0) + \tag{3.5}$$

$$+ \frac{t_{n+1} - t_\nu}{t_{n+1} - t_0} f_{a_1 + a_2 Z_{1,n} + \ldots + a_\nu Z_{\nu-1,n} + a_{\nu+1} Z_{\nu,n} + \ldots + a_{n+1} Z_{n,n}}(x - t_0) .$$

**PROOF.**

Follows by Lemma 3.1 and the fact that a B-spline on the grid $\underline{x}$ is a non-negative linear combination of B-splines on the refined grid $\underline{t}$ (see C. de Boor (1984), §2).

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

| REPORT DOCUMENTATION PAGE | | READ INSTRUCTIONS BEFORE COMPLETING FORM |
|---|---|---|
| 1. REPORT NUMBER #2817 | 2. GOVT ACCESSION NO. AD-A158 151 | 3. RECIPIENT'S CATALOG NUMBER |
| 4. TITLE (and Subtitle) B-Splines and Linear Combinations of Uniform Order Statistics | | 5. TYPE OF REPORT & PERIOD COVERED Summary Report - no specific reporting period |
| | | 6. PERFORMING ORG. REPORT NUMBER |
| 7. AUTHOR(s) Z. G. Ignatov and V. K. Kaishev | | 8. CONTRACT OR GRANT NUMBER(s) DAAG29-80-C-0041 |
| 9. PERFORMING ORGANIZATION NAME AND ADDRESS Mathematics Research Center, University of Wisconsin, 610 Walnut Street, Madison, Wisconsin 53706 | | 10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS Work Unit Number 4 - Statistics and Probability |
| 11. CONTROLLING OFFICE NAME AND ADDRESS U. S. Army Research Office, P.O. Box 12211, Research Triangle Park, North Carolina 27709 | | 12. REPORT DATE May 1985 |
| | | 13. NUMBER OF PAGES 23 |
| 14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office) | | 15. SECURITY CLASS. (of this report) UNCLASSIFIED |
| | | 15a. DECLASSIFICATION/DOWNGRADING SCHEDULE |

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

B-spline, discrete B-spline, linear combination of order statistics, Polya distribution

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)


Consider independent random variables $X_1, X_2, \ldots, X_n$, uniformly distributed on unequal intervals $(0, a_i)$, $a_i$ real, $i = 1,2,\ldots,n$. Let $S = X_1 + X_2 + \ldots + X_n$. The density of $S$ is shown to be a linear combination of $n!$ B-splines with constant coefficients. A probabilistic interpretation of the B-spline as the density of a linear combination of order statistics from the uniform distribution on $(0,1)$ is given. This interpretation makes it


DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

**ABSTRACT (continued)**

possible to establish recurrence relations for densities and moments of linear combinations of order statistics and to give asymptotic results both for B-splines and linear combinations as well. Examples and applications are also discussed.

END

FILMED

9-85

DTIC